DOCUMENT RESUME

ED 388 264 IR 017 413

AUTHOR Ki, W.; And Others
TITLE A Knowledge-based Multimedia System to Support the Teaching and Learning of Chinese Characters.
PUB DATE 94
NOTE 7p.; In: Educational Multimedia and Hypermedia, 1994. Proceedings of ED-MEDIA 94--World Conference on Educational Multimedia and Hypermedia (Vancouver, British Columbia, Canada, June 25-30, 1994); see IR 017 359.
PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Chinese; Computer Graphics; *Computer System Design; *Courseware; Educational Technology; Foreign Countries; *Ideography; *Language Acquisition; Language Patterns; Learning Activities; Learning Strategies; *Multimedia Materials

ABSTRACT

A current project is underway to develop a multimedia system that would support the teaching and learning of Chinese characters, as well as provide a platform for conducting research into the cognitive aspects of Chinese language acquisition. Although the number of commonly used Chinese characters amounts to thousands, there are many structural commonalities and regularities among the characters. One of the objectives in developing the system is to build a knowledge-base representing the structural features of and relationships among Chinese characters, where these structures can be made more explicit and explorable to the learner. The computer, with its interactive graphic ability, can provide a medium where character structures can be presented with flexible highlights directing learners' attention to various components; it can also provide a variety of activities where the learner can practice disassembling and assembling logographic components of Chinese characters. It is envisaged that through interactions with such environments, the learner would be able to develop a structural understanding about Chinese characters and acquire effective strategies for learning them. (Contains 17 references.) (AEF)

A Knowledge-based Multimedia System to Support the Teaching and Learning of Chinese Characters

W. KI, S.K. TSE, N. LAW
Department of Curriculum Studies
University of Hong Kong, Pokfulam Rd., Hong Kong
Email: hraskww@hkucc.hku.hk

F. LAU, K.H. PUN
Department of Computer Science
University of Hong Kong, Pokfulam Rd., Hong Kong

Abstract: This paper reports on a current project to develop a multi-media system that would support the teaching and learning of Chinese characters, as well as provide a platform for conducting research into the cognitive aspects of Chinese language acquisition. Although the number of Chinese characters commonly in use amounts to thousands, there are in fact many structural commonalities and regularities among the characters, which relate these characters into a highly connected system. One of the objectives in developing the present system is to build a knowledge-base to represent these structural features of and relationships among Chinese characters, so as to support the development of interactive learning environments where these underlying structures of Chinese characters can be made more explicit and explorable to the learner. It is envisaged that through interactions with such environments, the learner would be able to develop a structural understanding about Chinese characters and acquire effective strategies for learning them.

This project intends to help young first language learners, from pre-primary to early primary levels, to learn Chinese characters. In learning a character, the learner has to learn its written form, its pronunciation and its meaning. There are reasons to believe that before the school begins to teach these children the written form of Chinese, the children have already acquired, from their daily life, a rich mental lexicon for the language, in its audio form (Kwong 1992). So the major tasks to be achieved in teaching these characters are (1) to provide effective means for the learner to encode the written form of the characters, and (2) to help the learner to relate these written forms to sounds and meanings already existing in their mental lexicon (Huang & Liu 1978). In the following sections, the authors will examine how a computer system may provide support for this teaching and learning process.

Effective Learning of the Written Form of Chinese Characters

Ideographic Nature of Chinese Characters

Chinese characters are not just arbitrary aggregates of strokes. Many Chinese characters are pictographs standing for objects, or pictographs with certain markings added to indicate more abstract concepts (Shuowenjiezi, 100,Weiger 1911; Hung & Tseng 1981).

Research shows that at an early stage in the learning of Chinese characters, the learner tends to remember characters as distinct pictures (Chuang 1975). In one earlier project on software for learning Chinese characters carried out by the team (Lam 1993), computer animation was used to relate the written form of some Chinese characters to their pictorial origins. Preliminary evaluation indicated that such presentation is motivating as well as effective in helping young learners to remember the written form as well as the meanings of the characters.

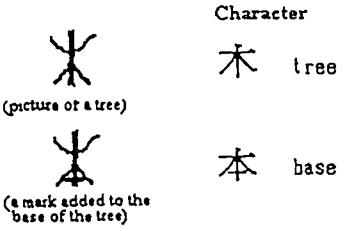

Figure 1. Chinese characters as pictorial and indicative symbols

Chinese Characters as Structured Compositions

Most Chinese characters are composite logographic forms (Shuwenjiezi, 100). In these "compound characters", two or more major components are combined to fill an imaginary square block. The methods of composition are highly regular. The components are reusable in producing different characters, and many of them are themselves characters. Some basic methods of composition are shown in figure 2, each illustrated with a compound character of that format.

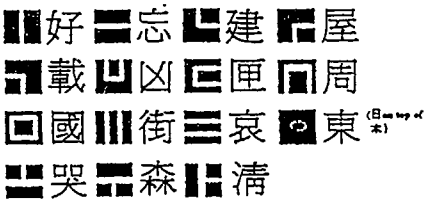

Figure 2. Chinese characters as structured compositions

In general, the components contribute to the sound or the meaning of the compound character they form (Ann 1982). In some cases, the components are themselves characters, and the meanings of these simple characters add together to form the meaning of the compound character.

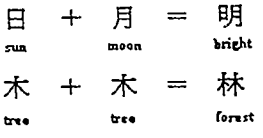

Figure 3. Characters representing an aggregate of meanings contributed by its components

In other cases, one component represents the meaning while the other component the sound of the compound character. This is the so called pictophonetic composition. Over 70% of Chinese characters are compositions of this sort.

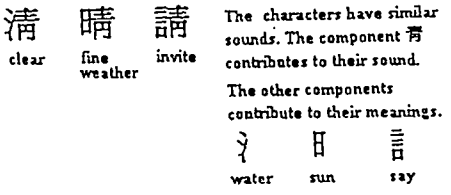

Figure 4. Examples of characters formed by pictophonetic composition

If one further examines these components, one will notice that they can in fact be further decomposed into simpler reusable stroke patterns. These we shall refer to as "graphemes". Although this decomposition to the graphemic level often does not add knowledge about the meaning or sound of the character, familiarity with the set of commonly used graphemes do help people in their recognition and writing of the characters or components (Huang 1992).

In studies conducted by Shen Yeh and Tsao Chuan Yung (1963, 1965), it was found that there were two major developmental changes in the process of mastering Chinese characters by primary school children: the first one was found immediately after the first month of learning Chinese characters and the other was between the first and second grades. With the second transition, the ability to master unfamiliar characters grew rapidly.

According to reported experience of some local school teachers, learners at grade two begin to demonstrate some mental processes of graphemic analysis and synthesis. Having learnt the characters 兄 (brother) and 兒 (son), a learner may identify the reusable grapheme "⺅". When presented with a new character 插 (insert), the learner may be able to describe verbally the component 臿 as a 千 (thousand) on top of "a 兒 without the '⺅' ". This ability to generate new logographic constructs to chunk new information would help to reduce significantly the memory load required in encoding the form of new characters (Hung & Wang 1992). Furthermore, some learners can guess that the character means some action, because they know other characters like 打 (hit), 推 (push), 拉 (pull) also have the same component 扌 and they all mean some actions.

Using Computers to Facilitate Structural Perception and Understanding of Chinese Characters

Although research seems to show structural encoding to be an important process in the learning of Chinese characters, current practice in primary grades is to drill learners in the stroke-by-stroke writing of characters (CDC 1990). Much of the learner's attention is thus directed to the individual strokes, while the development of structural perception of Chinese characters is very much left to his/her own initiative.

In this project, we see great potentials for the computer to provide useful learning experience, beyond the conventional handwriting exercises, for the learner to develop a structural perception of Chinese characters. The computer, with its interactive graphic capability, can provide a medium where character structures can be presented with flexible highlights directing learners' attention to the various components. The computer can also provide a rich variety of activities where the learner can practise disassembling and assembling logographic components of Chinese characters.

More importantly, the computer can allow the learner to learn the written form of Chinese characters with more meanings and interconnections. As pointed out earlier, there are many regularities in the construction of Chinese characters and the rich information encoded in the written form of a Chinese character may include cues about its meaning and sound. Hypertext systems can be designed such that characters related through particular structural features or components can be made readily accessible for the learner to examine. Characters can be learned in rich relationship with one another and learners can traverse from one character to the next through a shared component. Learners may even explore, like linguists, the regularities in the structure of Chinese characters. Knowing about the components of a character becomes an exciting pathway for information access and concept formation. In exploring this hypertext, learners can construct their own theories about the structure of Chinese characters. Such an approach to the learning of Chinese characters will be referred to as the *structural constructivist approach.*

Although Chinese characters are ideographs and meanings are attached to individual characters, such meanings are often not too specific. Most nouns, verbs and adjectives in Chinese are in fact made up of two or more characters. Such linguistic units will be referred to here by their original name in Chinese: *ci* (詞). In many instances, *ci* seem to correspond more to words in English, while Chinese characters to monosyllabic morphemes (Hoosain 1992, Huang & Liu 1978). *Ci* ususally have meanings much more specific than characters. For example, the idea of a policeman, corresponding to the *ci* 警察, is concrete and commonplace in everyday life, while the character 察 carries a more general or abstract notion of to notice or to examine. A possible way for young learners to learn the meaning of a character would be to explore the meaning of a number of commonly used *ci* containing the character and make their own induction. Here the hypertext capability of computer can again play an important role.

A Knowledge-based Multimedia System to Support the Learning of Chinese Characters

This section describes the design of the system we are currently developing. It is intended to be a flexible and extensible platform for the development of computer-aided-learning (CAL) software for Chinese character learning, as well as research into cognitive aspects of Chinese language learning. The system is designed to support, in particular, teaching and learning in the structural constructivist approach.

Conceptually, the system consists of five levels. The bottom layer of the system is a knowledge-base about structures, relationships, and meanings of Chinese characters. The second level is a database which records the frequency, extent of use and competence of a student in relation to the different elements in the knowledge-base. The third level is a set of application software providing access to the Chinese character knowledge-base (first level) and the student progress database (second level) for use by teachers and students. The fourth level is a set of courseware generation tools for use by teachers to generate customized CAL software. The fifth level is a set of CAL materials developed by tools in the lower levels for teaching various topics about Chinese characters.

The System Backbone: The Chinese Character Knowledge-base and Its Basic Functions

The Chinese character knowledge-base is built to represent the structural features and relationships among Chinese characters. It consists of two modules, one responsible for the logographic information of characters, and the other responsible for the meanings and sounds of characters.

The logographic information module contains a database of logograms, including characters, components, graphemes and strokes. It records how each logogram is stepwise-decomposable into other simpler logograms, and the composition method. In the case of pictophonetic composition, it also makes note of the specific subcomponents contributing respectively to the sound and the meaning of the character. Pictorial origins of characters are stored as graphic animation where appropriate.

The information stored will support the following operations:

(a) to give a stroke-by-stroke calligraphic display of the character.

(b) to provide a stroke ordering exercise where the learner can choose the correct stroke sequence in a character with continuous feedback from the computer.

(c) to display the character, highlighting in turn the different components and subcomponents.

(d) to support jigsaw-like exercises where the learner is asked to assemble characters from their components.

(e) partial matching to search for a character based on partial information about its structures and components

"Find all characters that have this structure."

This will return characters like:

辦,辨,辮,辯,瓣,..

Figure 5. Example of a search based on partial information on the written form

(f) to present the pictorial origins of some characters. The presentation starts with a realistic picture, and gradually transforms to a skeleton, and then to the written form of the character. The method is similar to that used in Lam (1993)

In *the sound and meaning module*, each character node is linked to a set of *ci* illustrating the meaning of the character. A voice file and a phoneme are stored for the pronunciation of each character in the knowledge-base. For the special cases where a character is pronounced differently when it is used in different senses, a number of voice files and phonemes will be stored. For some of the *ci*, additional resources are stored including images and voice files of sample sentences. Further, for each character and *ci*, there is information about its meaning category and difficulty level.

Such information is to support the following functions in this module:

(a) to display related *ci* to illustrate the meaning(s) of the character.

(b) to read aloud the characters and *ci* in the knowledge-base.

(c) to read aloud example sentences or display pictures for a particular *ci*.

(d) to search for characters that are pronounced identically.

(e) to search for characters based on partial information on phonemes.

(f) to search for *ci* according to their meaning categories, and difficulty levels.

To start with, the system will contain information for 500 Chinese characters selected from those specified in the school curriculum for young first language learners (CDC 84, 90). According to our analysis, we envisage that within these 500 characters, the structure and basic graphemes necessary for the learning of most Chinese characters would have been covered. We hope that by learning these 500 characters in the way supported by the system, the mental "tool-kit and library" can be developed so that the learner will be able to learn other Chinese characters proficiently (Shiu & Lau 1982, Tseng & Wang 1983).

The Other Levels of The System

The second level - the student progress database. The limitation of any teaching/learning material that does not take the learner's past experience and competence into account is widely recognized. The objective here is to set up a database that records for each student his/her usage profile and achieved competence as demonstrated by his/her interactions with the system. Data will be automatically logged into this database when the learner uses the various modules at the higher levels. This database can be used to provide feedback to students, to allow teachers to monitor the progress of students, and for use by the software generators at level 4 to generate appropriate CAL software.

The third level - database tools. This is a set of application software that allows both teachers and students to modify, customize and access the Chinese character knowledge-base and the student progress database.

For the teacher, there are four types of application software:

1. The knowledge-base construction tool - This is a user-friendly interface provided to allow modification and extension of the knowledge-base. Teachers may add in new characters, structures, sense nodes,*ci*, together with voice and image resources, and even new meaning categories. This is seen as an extremely important feature. It means that teaching of Chinese characters through the system need not be confined by the exact contents of the knowledge-base as provided. Teachers can change it and such changes can automatically be reflected in all the applications built on the knowledge-base.
2. The teacher query tool - This is a tool provided to teachers to query the knowledge-base for lesson planning and courseware development purposes.
3. The teacher customization tool - As with printed dictionaries and thesauruses, different versions will be required for different levels of users. The customization tool allows the teacher to index the knowledge-base into different levels of difficulty and allow student access to a designated portion only.
4. The reporting tool - This is for teachers to monitor the progress of individual students. It reports on the profile of usage: modes of usage, parts of the knowledge-base traversed, and performance in cases where the student has taken exercise or tests (level 4 and 5 applications) from the system.

For the student, there are three types of tools:

1. The student query tool - It may be used as a general reference tool by the student. The basic functions have been outlined in the previous section about the knowledge-base. Teachers may design specific learning tasks or research projects that require the use of the knowledge-base, or students may be allowed to call up this tool for help while they are working on other software on the system. .
2. The student customization tool - The purpose of this tool is to allow the student to create a personalized knowledge base of Chinese characters and *ci* that he/she wants to retain for easy access in future. Furthermore, the student may construct his/her own data-structure for organizing language elements (which in itself offers great potential for research into cognitive features in Chinese language learning).
3. The reporting tool - This is similar to the teacher reporting tool. It allows the student to find out how much he/she has learnt and mastered for self-monitoring purposes.

The fourth level - CAL generation tools. This is a set of software generation tools that can produce courseware in specific formats. The teacher just has to choose the courseware type, and specify the required language elements to be used, how he/she would like to draw reference from the student progress database, and how the student progress database should be updated according to the student's performance. The courseware will be automatically generated. One would envisage instruction like the following would be sufficient to generate a CAL program: "Generate a program on assembling characters with two components, in the form of a cat-catch--mouse game, using the set of characters in the knowledge-base that has been visited by the learner. Add one point to the learner's structured-perception index to a character if he/she succeeds in assembling the character."

The fifth level - CAL materials. The fifth level is a set of example courseware and self-learning materials, both computer-based and non-computer based, developed using the tools provided at the lower levels. It illustrates how the generic tools in level 4 can be use to achieve specific learning objectives. It includes learning materials that introduce new characters and *ci*, materials that focus on particular topics about Chinese characters or particular language skills, games for enhancing interest or for reinforcement, and tests. While these courseware are being used on the computer, the system will be able to update the student progress database at the second level of the system. Also, the learner can have the option of moving from the courseware to the student query tool for reference or help.

Conclusion

This paper reports on the considerations made in the development of a multi-media system that would support the teaching and learning of Chinese characters for young first language learners. The logographic and morphemic nature of Chinese characters are discussed, and a structural constructive model of learning is described. The knowledge-base is currently under construction and the application programs are now in the process of detail design. It is envisaged that the evaluation of such a system in classrooms is going to provide useful feedback on the various theoretical assumptions about the cognitive processes in Chinese language learning discussed in the paper.

References

Ann T.K. (1982), *Cracking the Chinese Puzzles*, H.K.:Stockflows Co., Ltd.

Atkinson,R.C. (1975) Mnemonics techniques in second-language learning. In *American Psychologist*, Vol. 30, pp. 821-828.

Chuang C.J. (1975), The function of imagery in learning of Chinese language, *Acta Psychologica Taiwanica*, Vol. pp.145-150.

Hong Kong Curriculum Development Committee (CDC) (1984), *Guide to the Kindergarten curriculum*, Hong Kong Education Dept.

-- (1990), *Chinese language syllabus for primary schools*, Hong Kong Education Dept.

Hoosain, R. (1992), Psychological reality of the word in Chinese. In*Language processing in Chinese*, H.C. Chen & O.J.L. Tzeng (eds). Amsterdam: Elsevier Science Publishers

Huang, J.T. & Liu I.M.(1978), Paired-associate learning proficiency as a function of frequency count, meaningfulness, and imagery value in Chinese two-character ideograms. *Acta Psychologica Taiwanica*,1 67, 5-17.

Huang J.T. & Wang M.Y.(1992) From unit to gestalt: perceptual dynamics in recognizing Chinese characters. In *Language processing in Chinese*, H.C. Chen & O.J.L. Tzeng (eds). Amsterdam: Elsevier Science Publishers

Hung, D.L., & Tseng, O.J.L. (1981) Orthographic variations and visual information processing. *Psychological Bulletin*, 90, 377-414.

Kwong, S.M. (1990) *The syntactic development of Cantonese-speaking preschool children.* M.Ed. Thesis, University of Hong Kong.

Lam,H.C., Pun,K.H., Leung,S.T., Tse,S.K. & Ki,W.W. (1993), Computer Assisted Learning for Learning Chinese Characters. In *Communications of COLIPS, an international journal of the Chinese and Oriental Languages Processing Society*, Vol. 3, No.1.

Liu Y.T. (1965), Experimental studies on the effects of pictorial material and writing in learning characters, *Acta Psychological Sinicia*, Vol. 3.

Paivio,A. & Casapo,K. (1973) Picture superiority in free recall: imagery or dual coding? In*Cognitive Psychology*, Vol. 5, pp 176-206.

Shiu P.K. & Lau S. (1982), The association pattern of Chinese characters of Hong Kong kindergarten children. In *Psychological Studies of Chinese Language*, S.R.Kao & C.M.Chen (eds.) pp.103-122, Taiwen:Wenhe.

Shuowenjiezi (A.D.100), *The analysis of Chinese characters.* H.K.:Chung Wah (reprint)

Tsao C.Y. & Shen Y. (1965), Developmental studies on the recognition and generalization of Chinese characters in primary school children. *Acta Psychologica Sinica*, Vol 1 & 2.

Tzeng, O.J.L. & Wang W.S.Y. (1983), The first two R's. *American Scientist*, 71, pp. 238-243.

Weiger, L (1915), *Chinese characters, their origin, etymology and history.* Catholic Mission Pub.

Acknowledgment

The authors would like to acknowledge the contribution of the other team members in their contribution to the software development project: M.K.Lau, H.C.Lam, S.S.Lam, Y.T.Lee, W.H.Chau & W.K.Tam The project is sponsored by the Hong Kong University and Polytechnic Grant Committee research grant, 1991. (hku 33/91)